

Free Energy Calculation Methods: A Theoretical and Empirical Comparison of Numerical Errors and a New Method for Qualitative Estimates of Free Energy Changes

RANDALL J. RADMER, PETER A. KOLLMAN

Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143-0446

Received 12 February 1996; accepted 30 September 1996

ABSTRACT: We present a comparison of four free energy calculation methods: thermodynamic integration (TI); traditional free energy perturbation (FEP); Bennett's acceptance ratio method (IPS); and a method that is related to an implementation of the WHAM method (CRS). The theoretical bases of the methods are first described, then calculations of the solvation free energies of methane and ethane are performed to determine the magnitude of the errors for the different methods. We find that the methods give similar errors when many intermediate states (windows) are used, but the IPS and CRS methods give smaller errors than the TI and FEP methods when no intermediate states are used. We also present a new procedure (based on the CRS method) that uses coordinates from simulations of a set of solutes to calculate the solvation free energies of additional solutes for which no simulations were performed. Solvation free energies for nine solutes (methanol, dimethylether, methylamine, methylammonium, dimethylamine, fluoromethane, difluoromethane, trifluoromethane, and tetrafluoromethane) are estimated based only on simulations of set of small hydrophobic solutes (including methane, ethane, and

Correspondence to: P. A. Kollman; e-mail: pak@cgl.ucsf.edu

Contract grant sponsor: National Institutes of Health, contract grant number: GM-39552

Contract grant sponsor: National Science Foundation, Contract grant number: CHE-94-13472

propane). These estimates can be surprisingly accurate and appear to be useful for making rapid estimates of solvation free energies. © 1997 by John Wiley & Sons, Inc. *J Comput Chem* **18**: 902–919, 1997

Introduction

Free energies calculated using molecular dynamics (MD) or Monte Carlo (MC) simulations are commonly used to obtain quantities such as partition coefficients, relative stabilities of macromolecules, and free energies of binding of small organic molecules to proteins and nucleic acids.^{1–6} Although frequently accurate, these free energies are only estimates of the correct free energies because two approximations are typically employed. The first is that empirically derived, classical energy functions are used to model interaction energies. The second approximation is that numerical methods are used to calculate ratios of configuration integrals that cannot be solved analytically. The numerical calculations are done using a set of atomic coordinates (generated by MD or MC simulations) as a representative sample of all possible configurations of the system. Because only a finite number of coordinate sets are generated, the calculated free energy is only an estimate of the correct free energy. Not surprisingly, more accurate estimates are obtained when longer simulations are used, but the minimum simulation time needed depends on the system being examined and cannot be determined before the simulation is performed. Pearlman⁷ has reported that, for some calculations involving simple perturbations, at least 700 ps of sampling was required for converged solvation free energies. It is reasonable to assume that more complex perturbations will require longer simulations.

Because of the amount of computer time needed to calculate free energies, the most efficient method should be used. In this article, we examine the errors resulting from the use of numerical methods in an effort to determine when the use of these methods is appropriate. First, we give a theoretical description of some of the different methods and discuss the expected errors for each of them. Then, we perform an empirical comparison of the four free energy calculation methods by determining

the errors in the calculated solvation free energies of a set of solutes. The four methods are: thermodynamic integration (TI); traditional free energy perturbation (FEP); Bennett's acceptance ratio method,¹ referred to by us as the intermediate perturbed state (IPS) method; and a method we refer to as the composite reference state (CRS) method that is related to an implementation of the WHAM method presented by Kumar et al.² It is worth noting that we are comparing methods of calculating free energies from a given simulation; we are not comparing simulation protocols that may also affect the accuracy of the free energy estimates.

When calculating relative free energies using the methods examined, it is traditional, and generally appropriate, to base the calculations on simulations of all relevant molecules. However, under conditions we will discuss, it is sometimes possible to estimate relative free energies for a set of molecules based on simulations of only a subset of molecules. We present preliminary results showing that, based on simulations of one set of solutes, it is possible to estimate the relative solvation free energies of nine additional solutes with reasonable accuracy. Although less accurate than calculations in which all molecules are simulated, this procedure appears to give qualitative free energy estimates for a wide variety of molecules without requiring that simulations be performed on each of them.

Theoretical Background

This section describes two methods for calculating relative free energies, each of which is based on an exact expression. As these cannot be solved analytically for most relevant systems, numerical methods are often employed. The resulting errors, particularly those resulting from insufficient sampling of the relevant systems, will be examined in greater detail and are the focus of this study.

The first method we will examine is thermodynamic integration,^{8–10} based on the following

equation:

$$\Delta G_{10} = \int_0^1 \frac{dG_\lambda}{d\lambda} d\lambda = \int_0^1 \left\langle \frac{d\mathcal{H}_\lambda}{d\lambda} \right\rangle_\lambda d\lambda \quad (1)$$

where $\Delta G_{10} = G_1 - G_0$ is the difference in Gibbs free energies between states 1 and 0. G_λ is the free energy of the state defined by a coupling parameter, λ , where λ can take on values from zero to unity. The total energy (Hamiltonian) is given by \mathcal{H}_λ , where \mathcal{H}_0 and \mathcal{H}_1 are the energies at states 0 and 1. The angle brackets, $\langle \dots \rangle_\lambda$, indicate a constant pressure, constant temperature ensemble average performed on the state with potential energy defined by λ .

In practice, the kinetic component of the free energy is not included in the numerical calculations described in what follows because it can generally be calculated analytically,⁹ or it can be made to cancel if thermodynamic cycles are used.¹¹ To keep the notation as simple as possible, we will use G to indicate the free energies calculated, but it should be kept in mind that these would be described more accurately as excess free energies.

The free energies are calculated by assuming that an infinite time average is equal to an ensemble average (by the ergodic hypothesis), a finite number of samples are used to estimate the infinite time average, and the integration is done numerically. If the trapezoid rule is used for the integration over λ , the free energy is estimated by:

$$\Delta G_{10} \approx \sum_{k=1}^N \left(\frac{1}{n_\lambda} \sum_{i=1}^{n_\lambda} \frac{d\mathcal{H}_\lambda(\mathbf{r}_{i,k})}{d\lambda} \right) \Delta\lambda \quad (2)$$

where N is the total number of simulations performed at different states (defined by λ); n_λ is the number of samples from each state; \mathcal{H}_λ is the potential energy; and the coordinates, $\mathbf{r}_{i,k}$, are chosen with a Boltzmann probability ($e^{[\mathcal{H}_\lambda(\mathbf{r}_{i,k}) - G_\lambda]/RT}$) by performing an MD or MC simulation at each λ (the subscript i indicates sample number and the k indicates simulation number).

The second method examined is free energy perturbation,⁸⁻¹⁰ using the following equation:

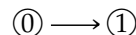
$$\Delta G_{10} = -RT \ln \langle e^{-\Delta\mathcal{H}_{10}/RT} \rangle_0 \quad (3)$$

where R is the gas constant, T is the absolute temperature, $\Delta\mathcal{H}_{10}$ is the energy difference between states 1 and 0, and $\langle \dots \rangle_0$ indicates an ensemble average at state 0.

In this article, we are primarily interested in different methods for solving eq. (3) numerically. These will be described in greater detail below.

TRADITIONAL FREE ENERGY PERTURBATION WITH ONE WINDOW

When eq. (3) is used to calculate relative free energies, state 0 (and in general any state on which a simulation is performed) is referred to as the *reference* state and state 1 (and any state that is not simulated) is referred to as the *perturbed* state. This relationship will be represented schematically as follows:



In practice, the kinetic component of the free energy is not included in the numerical calculations for reasons previously described. The relative free energy is estimated by first assuming that infinite time averages are equal to ensemble averages and, second, by using finite time averages to approximate infinite time averages. This gives:

$$\Delta G_{10} \approx \Delta g_{10} = -RT \ln \frac{1}{n_0} \sum_{i=1}^{n_0} e^{-\Delta\mathcal{H}_{10}(\mathbf{r}_{i,0})/RT} \quad (4)$$

where Δg_{10} indicates a finite time estimate of ΔG_{10} ; n_0 is the number of samples on state 0; and coordinates, $\mathbf{r}_{i,0}$, are chosen with a Boltzmann probability by performing a MD or MC simulation on state 0.

Errors in the free energy estimates given by eq. (4) can be characterized by examining an expression for the root mean square (RMS) error. It can be shown (see Appendix) that for a *large number* of *independent* samples, an approximate expression for the RMS error in the estimate of ΔG_{10} is:

$$RMS(\Delta g_{10} - \Delta G_{10}) \approx RT \left(\frac{1}{n_0} \int \frac{\rho_1^2(\mathbf{r})}{\rho_0(\mathbf{r})} d\mathbf{r} - \frac{1}{n_0} \right)^{\frac{1}{2}} \quad (5)$$

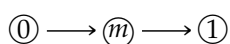
where $\rho(\mathbf{r}) = e^{-[\mathcal{H}(\mathbf{r}) - G]/RT}$ is the Boltzmann probability.

Two points regarding eq. (5) are worth noting. The first is the well-known fact that the RMS error is proportional to the reciprocal of the square root of the number of independent samples (RMS error $\propto \frac{1}{\sqrt{n}}$); thus, increasing sampling decreases the expected error. The second point concerns how the relative probabilities of states 0 and 1 affects the error. Because the ratio of $\rho_1^2(\mathbf{r})$ and $\rho_0(\mathbf{r})$ in eq. (5) must remain small to keep the RMS error small,

the probability of each conformation, \mathbf{r} , sampled by the reference state, $\rho_0(\mathbf{r})$, should never approach zero more quickly than the probability for the perturbed state, $\rho_1(\mathbf{r})$. Stated more simply (but with less accuracy), it should be possible to sample important conformations of the perturbed state during a simulation of the reference state. This leads to the common view that two states must be similar for the free energy perturbation method to give good estimates of their relative free energy. In fact, this is only necessary if both states are to be simulated, and the relative free energy of the other state calculated. As we will discuss later, it is often advantageous to simulate only one state, while ensuring that important perturbed state conformations are reasonably likely to be sampled. This requires only that the accessible perturbed state conformations be a subset of the accessible reference state conformations.

TRADITIONAL FREE ENERGY PERTURBATION WITH MULTIPLE WINDOWS

One way to decrease the RMS error is to add additional states between the two states of interest. The calculation then involves combining results from a series of smaller calculations, where each calculation is referred to as a window. For simplicity we will discuss the use of only one intermediate state, state m . A calculation from state 0 to state 1 can be expanded into two calculations: the first from state 0 to state m , and the second from state m to state 1:



The estimate of the total change in free energy is [see eq. (4)] is given by:

$$\begin{aligned} \Delta g_{10} &= \Delta g_{1m} + \Delta g_{m0} \\ &= -RT \ln \frac{1}{n_m} \sum_{i=1}^{n_m} e^{-\Delta \mathcal{V}_{1m}(\mathbf{r}_{i,m})/RT} \\ &\quad - RT \ln \frac{1}{n_0} \sum_{i=1}^{n_0} e^{-\Delta \mathcal{V}_{m0}(\mathbf{r}_{i,0})/RT} \end{aligned} \quad (6)$$

where $\mathbf{r}_{i,0}$ are coordinates generated by performing a simulation on state 0, $\mathbf{r}_{i,m}$ are coordinates generated by performing a simulation on state m , and n_0 and n_m are the number of samples on states 0 and m , respectively.

Samples on states 0 and m are independent, so the RMS error can be given as follows [see eq. (5)]:

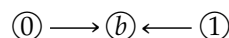
$$\begin{aligned} \text{RMS}(\Delta g_{10} - \Delta G_{10}) \\ \approx RT \left(\frac{1}{n_0} \int \frac{\rho_m^2(\mathbf{r})}{\rho_0(\mathbf{r})} d\mathbf{r} \right. \\ \left. - \frac{1}{n_0} + \frac{1}{n_m} \int \frac{\rho_1^2(\mathbf{r})}{\rho_m(\mathbf{r})} d\mathbf{r} - \frac{1}{n_m} \right)^{\frac{1}{2}} \end{aligned} \quad (7)$$

This equation is useful because it makes it possible to compare the errors in this method versus a calculation with no intermediates; if the energy of state m is defined as the average of the energies of state 0 and state 1 (and equal total sampling is done), then subtracting the square of eq. (7) from the square of eq. (5) always gives a positive number, indicating that this method does give a smaller expected error than methods using no intermediate states.

The definition of state m that minimizes eq. (7) could be found in the same manner as will be done in the next two sections, but it does not address the essential problem with this method. The difficulty is that state m acts as both a reference state, which should be able to sample all important configurations of state 1, and as a perturbed state, which should be completely accessible from sampled configurations of state 0. Arranging the calculation so that the intermediate states are only reference states or perturbed states eliminates this problem and will be described in what follows.

INTERMEDIATE PERTURBED STATE (IPS) METHOD

Sampling could be done from state 0 to a new intermediate perturbed state (state b) and, independently, from state 1 to the same intermediate state. The total free energy change can be found by subtracting these two free energy differences. An optimal state b will be defined such that the expected RMS error in the estimated free energy difference between states 0 and 1 is minimized. Although we will present it in a different form, this is formally identical to Bennett's acceptance ratio method¹:



The total change in free energy is [see eq. (4)]:

$$\begin{aligned}\Delta g_{10} &= \Delta g_{b0} - \Delta g_{b1} \\ &= -RT \ln \frac{\frac{1}{n_0} \sum_{i=1}^{n_0} e^{-\Delta \mathcal{Z}_{b0}(\mathbf{r}_{i,0})/RT}}{\frac{1}{n_1} \sum_{i=1}^{n_1} e^{-\Delta \mathcal{Z}_{b1}(\mathbf{r}_{i,1})/RT}} \quad (8)\end{aligned}$$

and the RMS error (see Appendix) is given by:

$$\begin{aligned}RMS(\Delta g_{10} - \Delta G_{10}) &\approx RT \left(\int \left[\frac{\rho_b^2(\mathbf{r})}{n_0 \rho_0(\mathbf{r})} + \frac{\rho_b^2(\mathbf{r})}{n_1 \rho_1(\mathbf{r})} \right] \right. \\ &\quad \left. \times d\mathbf{r} - \frac{1}{n_0} - \frac{1}{n_1} \right)^{\frac{1}{2}}\end{aligned}$$

The definition of state b that minimizes this equation (see Appendix) is:

$$\rho_{b,opt}(\mathbf{r}) = \lambda_b \frac{n_0 \rho_0(\mathbf{r}) n_1 \rho_1(\mathbf{r})}{n_0 \rho_0(\mathbf{r}) + n_1 \rho_1(\mathbf{r})} \quad (9)$$

where λ_b is the normalization constant. Equation 9 indicates that the accessible configurations of the optimal perturbed state are essentially the configurations accessible to *both* of the reference state, ensuring that accessible perturbed state conformations can be sampled during simulations of either reference state.

The relative free energy, ΔG_{10} , is estimated by solving for $\mathcal{Z}_b(\mathbf{r})$ in eq. (9) [using $\rho_{b,opt}(\mathbf{r}) = e^{-[\mathcal{Z}_b(\mathbf{r}) - G_b]/RT}$], substituting this into eq. (8), and rearranging and canceling terms:

$$\Delta g_{10} = -RT \ln \frac{\frac{1}{n_0} \sum_{i=1}^{n_0} \frac{e^{-\mathcal{Z}_1(\mathbf{r}_{i,0})/RT}}{q_b(\mathbf{r}_{i,0})}}{\frac{1}{n_1} \sum_{i=1}^{n_1} \frac{e^{-\mathcal{Z}_0(\mathbf{r}_{i,1})/RT}}{q_b(\mathbf{r}_{i,1})}} \quad (10)$$

where:

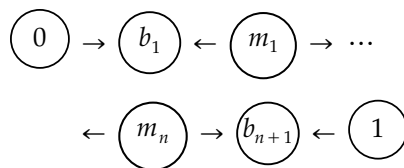
$$q_b(\mathbf{r}) = n_0 e^{-[\mathcal{Z}_0(\mathbf{r}) + \frac{1}{2}\Delta G_{10}]/RT} + n_1 e^{-[\mathcal{Z}_1(\mathbf{r}) - \frac{1}{2}\Delta G_{10}]/RT} \quad (11)$$

It is generally advantageous to rearrange these equations so that the exponentials contain relative, rather than absolute, energies to prevent numerical overflow.

To solve eqs. (10) and (11) for Δg_{10} , the value of ΔG_{10} is needed. As this is obviously not known, an estimate of ΔG_{10} must be used (this is still correct

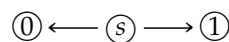
in the large sample limit, although it is not optimal). Two simple ways of dealing with this are to set ΔG_{10} to zero (we call this the IPSO method) or to set ΔG_{10} equal to Δg_{10} . For the second case, the free energy differences can be found by substituting Δg_{10} for ΔG_{10} in eqs. (10) and (11), and solving for Δg_{10} (Numerical Recipes¹² gives efficient methods for finding roots of equations). A convenient method used by Bennett¹ consists of guessing an initial value of Δg_{10} , using this in eq. (11), and using the resulting $q_b(\mathbf{r}_i)$ to get an improved estimate of Δg_{10} using eq. (10). This process is repeated until Δg_{10} is self-consistent (we call this the IPSsc method). A minor disadvantage of this method is that it requires storing the potential energies found during the simulation and calculating the free energies after the simulations are completed.

To get accurate results, some overlap must exist between states 0 and 1. If the states are too dissimilar, additional simulations can be performed on appropriately chosen intermediate states and the calculation done for each of the smaller perturbations as described previously. This is illustrated below for n arbitrary intermediate states (m), resulting in $n + 1$ IPS states (b).



INTERMEDIATE REFERENCE STATE METHODS

A new intermediate reference state (state s) can be defined, making states 0 and 1 perturbed states. Although we use different notation, this is equivalent to using umbrella sampling to find the free energy difference (see Torrie and Valleau¹³):



The estimate change in free energy is [see eq. (4)]:

$$\begin{aligned}\Delta g_{10} &= \Delta g_{1s} - \Delta g_{0s} \\ &= -RT \ln \frac{\frac{1}{n_s} \sum_{i=1}^{n_s} e^{-\Delta \mathcal{Z}_{1s}(\mathbf{r}_{i,s})/RT}}{\frac{1}{n_s} \sum_{i=1}^{n_s} e^{-\Delta \mathcal{Z}_{0s}(\mathbf{r}_{i,s})/RT}} \quad (12)\end{aligned}$$

and the RMS error, is given by:

$$\begin{aligned} \text{RMS}(\Delta g_{10} - \Delta G_{10}) \\ \approx RT \left(\frac{1}{n_s} \int \frac{[\rho_0(\mathbf{r}) - \rho_1(\mathbf{r})]^2}{\rho_s(\mathbf{r})} d\mathbf{r} \right)^{\frac{1}{2}} \end{aligned}$$

(see Appendix). Unfortunately, the optimal reference state found using this equation is not useful. Although it minimizes the RMS error in the estimate of ΔG_{10} , it also allows the RMS error in estimates of the intermediate free energies (Δg_{1s} and Δg_{0s}) used in eq. 12 to approach infinity. So, in practice, Δg_{1s} and Δg_{0s} cannot be calculated if the simulation is performed at the optimal reference state.

In principle, any state that samples all accessible configurations of states 0 and 1 can be used, but this may be difficult to choose in practice. A computationally convenient definition of state s results from generating a set of coordinates from state 0 (by performing a simulation on state 0) and generating another set of coordinates from state 1 (by performing a simulation on state 1). The combination of the two sets of coordinates (in the large sample limit) can be considered to have resulted from a single simulation on a state with probability, $\rho_{s,ave}(\mathbf{r})$, equal to the weighted average of $\rho_0(\mathbf{r})$ and $\rho_1(\mathbf{r})$:

$$\rho_{s,ave}(\mathbf{r}) = \frac{n_0 \rho_0(\mathbf{r}) + n_1 \rho_1(\mathbf{r})}{n_0 + n_1} \quad (13)$$

This is used to estimate ΔG_{10} by solving for $\mathcal{V}_s(\mathbf{r})$ in eq. (13) [using $\rho_{s,ave}(\mathbf{r}) = e^{-[\mathcal{V}_s(\mathbf{r}) - G_s]/RT}$]. From eq. (12), equations giving g_{0s} and g_{1s} can be found:

$$\begin{aligned} \Delta g_{0s} = -RT \ln \frac{1}{n_0 + n_1} & \left(\sum_{i=1}^{n_0} \frac{e^{-\mathcal{V}_0(\mathbf{r}_{i,0})/RT}}{e^{-\mathcal{V}_s(\mathbf{r}_{i,0})/RT}} \right. \\ & \left. + \sum_{i=1}^{n_1} \frac{e^{-\mathcal{V}_0(\mathbf{r}_{i,1})/RT}}{e^{-\mathcal{V}_s(\mathbf{r}_{i,1})/RT}} \right) \end{aligned}$$

and:

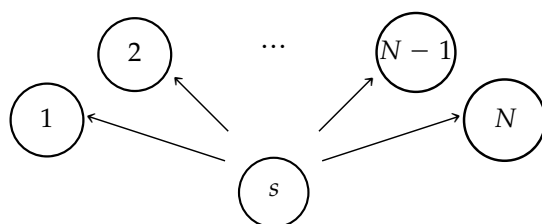
$$\begin{aligned} \Delta g_{1s} = -RT \ln \frac{1}{n_0 + n_1} & \left(\sum_{i=1}^{n_0} \frac{e^{-\mathcal{V}_1(\mathbf{r}_{i,0})/RT}}{e^{-\mathcal{V}_s(\mathbf{r}_{i,0})/RT}} \right. \\ & \left. + \sum_{i=1}^{n_1} \frac{e^{-\mathcal{V}_1(\mathbf{r}_{i,1})/RT}}{e^{-\mathcal{V}_s(\mathbf{r}_{i,1})/RT}} \right) \end{aligned}$$

where:

$$\begin{aligned} e^{-\mathcal{V}_s(\mathbf{r})/RT} \\ = \frac{n_0 e^{-[\mathcal{V}_0(\mathbf{r}) - \Delta G_{0s}]/RT} + n_1 e^{-[\mathcal{V}_1(\mathbf{r}) - \Delta G_{1s}]/RT}}{n_0 + n_1} \end{aligned}$$

Because ΔG_{0s} and ΔG_{1s} are not known, their values will be estimated using Δg_{0s} and Δg_{1s} . The relative free energies can then be determined self-consistently as was done with the IPSsc method [see eqs. (10) and (11)].

This methodology can be extended to more than two states by mixing additional states into $\rho_s(\mathbf{r})$, and performing all perturbations with respect to state s . For a total of N states, this can be represented as:



where state s is defined such that its Boltzmann probability is equal to the weighted average of the Boltzmann probabilities of all N states:

$$\rho_{s,ave}(\mathbf{r}) = \frac{1}{\sum_{k=1}^N n_k} \sum_{k=1}^N n_k \rho_k(\mathbf{r})$$

where n_k is the number of samples in the simulation of state k and $\rho_k(\mathbf{r})$ is the Boltzmann probability of state k . The set of N equations is represented by the following equation for the estimate of the free energy difference between state l and state s :

$$\Delta g_{ls} = -RT \ln \frac{1}{\sum_{k=1}^N n_k} \sum_{k=1}^N \sum_{i=1}^{n_k} \frac{e^{-\mathcal{V}_l(\mathbf{r}_{i,k})/RT}}{e^{-\mathcal{V}_s(\mathbf{r}_{i,k})/RT}} \quad (14)$$

where:

$$e^{-\mathcal{V}_s(\mathbf{r})/RT} = \frac{1}{\sum_{k=1}^N n_k} \sum_{k=1}^N n_k e^{-[\mathcal{V}_k(\mathbf{r}) - \Delta g_{ks}]/RT} \quad (15)$$

and l represents any state. [It is worth noting that eqs. (14) and (15) indicate that the energy of each state must be calculated for all coordinates generated, regardless of the state on which the coordinates were generated.] We will refer to this as the composite reference state (CRS) method because all of the perturbations are done with respect to one composite reference state (state s).

The energies in eqs. (14) and (15) can be found by periodically saving coordinates from each simu-

lation and then calculating the energy for all of the states for each saved coordinate. Although this obviously uses considerable disk space, it has the advantage of allowing any intermediate states to be used, and, additionally, it allows the coordinates to be used for additional free energy calculations after the simulation is completed. This is the method used for the H7 simulations we will describe.

The potential energy of intermediate states can often be defined as a linear combination of the potential energies of the two states of interest:

$$\mathcal{Z}_\lambda(\mathbf{r}) = (1 - \lambda)\mathcal{Z}_0(\mathbf{r}) + \lambda\mathcal{Z}_1(\mathbf{r}) \quad (16)$$

where λ is a coupling parameter that defines the states. In this case, coordinates do not have to be saved. Instead, simulations are done on all the states and the derivative of the potential energy with respect to λ is saved. Given $\frac{d\mathcal{Z}_\lambda(\mathbf{r})}{d\lambda}$ and λ for each coordinate of each simulation, the potential energies can be found with respect to the energy at $\lambda = 0$, and the relative free energies can be found using eqs. (14) and (15). This corresponds to an implementation of the WHAM method presented by Kumar et al.² [see eq. (21) in Kumar et al.²]. We use this method for the MM1, MM400, ME1, and ME400 simulations described in what follows.

When coordinates are saved, it is also possible to estimate the free energy of a state for which no simulations were done; in this case, state l in eqs. (14) and (15) does not correspond to any of the states 1 through N . For reasons previously discussed, care should be taken to assure that the composite reference state can sample all accessible configurations of the perturbed state.

Methodology

A total of five sets of simulations were performed such that relative free energies could be calculated for: (1) seven hydrophobic solutes (H7); (2) methane and methane using 1 window (MM1); (3) methane and methane using 400 windows (MM400); (4) methane and ethane using 1 window (ME1); (5) and methane and ethane using 400 windows (ME400).

In each case, the solute was simulated using AMBER 4.0,¹⁴ in a box of 207 TIP3P¹⁵ water molecules, at a constant temperature of 300 K and constant pressure of 1 bar using coupling con-

stants¹⁶ of 0.2 ps, and a time step of 1.0 fs. The SHAKE¹⁷ algorithm was used to constrain all bond-lengths to their equilibrium values and an 8-Å cutoff was used for the nonbonded interactions. Bond, angle, and torsional parameters are from the Weiner et al.¹⁸ force field. Only solute-solvent interaction energies were included in the calculation of the free energies.¹⁹

The H7 simulations were performed on a set of solutes with topology shown in Figure 1. Atomic parameters were chosen such that the solute would correspond to "nothing," methane, ethane, propane, a single generic atom (X), a diatomic molecule consisting of two generic atoms (X_2), or a triatomic molecule consisting of three generic atoms (X_3). The atomic charges are given in Table I and the van der Waals parameters are given in Tables II and III. At least 100 ps of equilibration followed by 500 ps of data collection was done for each solute. During the data collection, atomic coordinates were saved every 0.1 ps.

Three methods were used to determine the relative solvation free energies from the H7 simulations. The first was the FEP method using eq. (4) (this was done for both the forward and reverse directions). The second was the IPS method using eqs. (10) and (11) (solving for ΔG_{21} self-consistently), and the last was the CRS method, using eqs. (14) and (15), where the composite reference state consists of the seven solutes described previously. All calculations were done using only the saved coordinates. To determine the energy of the perturbed states, the coordinates were modified such that they would correspond to the correct bond-lengths as necessary. Resulting free energies are given in Table IV.

The MM1 and MM400 simulation sets were used to estimate the RMS error in the calculated

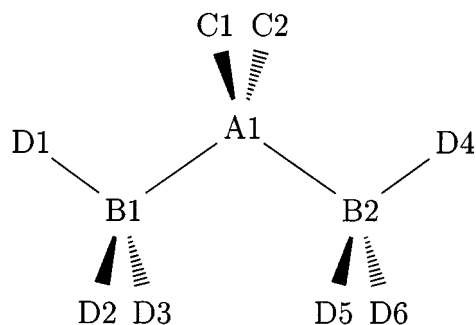


FIGURE 1. Topology for H7 solutes. See Table I for charges, and Tables II and III for van der Waal parameters.

TABLE I.
Atomic Charges Used for Each Solute (See Fig. 1).

Solute	A1	B1	B2	C1	C2	D1	D2, D3	D4–D6
Nothing	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Methane	−0.464	0.116	0.116	0.116	0.116	0.000	0.000	0.000
Ethane	−0.027	−0.027	0.009	0.009	0.009	0.009	0.009	0.000
Propane	0.296	−0.308	−0.308	−0.041	−0.041	0.067	0.067	0.067
X	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
X ₂	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
X ₃	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CH ₃ OH	0.196	−0.667	0.016	0.016	0.016	0.423	0.000	0.000
O(CH ₃) ₂	−0.344	−0.029	−0.029	0.067	0.067	0.067	0.067	0.067
CH ₃ NH ₂	0.307	−1.002	−0.017	−0.017	−0.017	0.000	0.373	0.000
CH ₃ NH ₃ ⁺	−0.038	−0.192	0.114	0.114	0.114	0.296	0.296	0.000
(CH ₃) ₂ NH	−0.691	−0.011	−0.011	0.373	0.000	0.057	0.057	0.057
CH ₃ F	0.011	−0.251	0.080	0.080	0.080	0.000	0.000	0.000
CH ₂ F ₂	0.390	−0.246	−0.246	0.051	0.051	0.000	0.000	0.000
CHF ₃	0.644	−0.230	−0.230	−0.230	0.046	0.000	0.000	0.000
CH ₄	0.756	−0.189	−0.189	−0.189	−0.189	0.000	0.000	0.000

solvation free energy difference between two methane solutes (see Fig. 2). Because of numerical errors, the calculated free energy difference will not, in general, be zero. The MM1 simulations involved 18 consecutive, 200.5-ps simulations of methane(1) and 18 consecutive 200.5-ps simulations of methane(2). Individual methane(1) simulations were paired with individual methane(2) simulations to give 18 sets of 401-ps simulations. No parameters were changed during these simula-

tions, so no equilibration beyond the initial equilibration was done. The MM400 simulations also involved 18 sets of 401-ps simulations. However, each simulation involved 401 windows, 1 ps in length. During each simulation the solute parameters were changed such that the interaction energy was a linear combination of methane(1) and methane(2) [see eq. (16) and Fig. 2], resulting in the consecutive simulation of methane(1), 399 intermediate solutes, and methane(2). The initial

TABLE II.
Atom van der Waals Parameters by Solute (See Fig. 1 and Table III).

Solute	A1	B1	B2	C1	C2	D1	D2, D3	D4–D6
Nothing	DH	DC	DC	DH	DH	DH	DH	DH
Methane	CT	HC	HC	HC	HC	DH	DH	DH
Ethane	CT	CT	HC	HC	HC	HC	HC	DH
Propane	CT	CT	CT	HC	HC	HC	HC	HC
X	X	DC	DC	DH	DH	DH	DH	DH
X ₂	X	X	DC	DH	DH	DH	DH	DH
X ₃	X	X	X	DH	DH	DH	DH	DH
CH ₃ OH	CT	OH	HC	HC	HC	HO	DH	DH
O(CH ₃) ₂	OS	CT	CT	H1	H1	H1	H1	H1
CH ₃ NH ₂	CT	NX	H1	H1	H1	DH	HX	DH
CH ₃ NH ₃ ⁺	CT	N3	HP	HP	HP	H	H	DH
(CH ₃) ₂ NH	NX	CT	CT	HX	DH	H1	H1	H1
CH ₃ F	CT	F	HF	HF	HF	DH	DH	DH
CH ₂ F ₂	CT	F	F	HF	HF	DH	DH	DH
CHF ₃	CT	F	F	F	HT	DH	DH	DH
CH ₄	CT	F	F	F	F	DH	DH	DH

TABLE III.
van der Waals Parameters by Atom Type
(See Table II).

Atom type	$R^*(\text{\AA})$	ϵ (kcal / mol)
CT	1.9082	0.1094
DC	1.9082	0.0000
DH	1.4872	0.0000
F	1.7500	0.0610
H	0.6000	0.0157
H1	1.3870	0.0157
HC	1.4872	0.0157
HF	1.2900	0.0150
HO	0.0000	0.0000
HP	1.1000	0.0157
HT	1.2100	0.0150
HX	0.6890	0.0157
OH	1.7210	0.2104
OS	1.6837	0.1700
N3	1.8240	0.1700
NX	1.8750	0.1700
X	1.6000	0.1000

structure for each simulation was taken from the final state of the previous simulation followed by 12 ps of equilibration. For each of the 1.0-ps simulations, data were generated using only the last 0.5 ps, allowing the system time to equilibrate. The MM1 and MM400 simulations used about the same total CPU time, but half of the MM400 simulation time was devoted to equilibration and half to data collection. No equilibration was done for the MM1 simulation, so only data collection was done. The derivative of energy with respect to λ ($\frac{d\mathcal{V}_\lambda(\mathbf{r})}{d\lambda}$) was stored at each step so that relative energies could be calculated [see eq. (16) and associated text].

Four free energy calculation methods were used to calculate the free energy difference from the data generated. The first was the TI method using eq. (2) and the second was the FEP method using eq. (4), which was done for both the forward (FEPf) and reverse (FEPr) directions. The combined (FEPc) value is the average of the forward and reverse calculations. The third method was the IPS method, using eqs. (10) and (11). The intermediate perturbed state was defined in two ways: by assigning a value of zero to ΔG_{21} on the right-hand side of eq. (11) (IPSO); or by solving for ΔG_{21} self-consistently (IPSsc) as described. The fourth method was the CRS method, using eqs. (14) and (15), where the composite reference state consists

of methane(1) and methane(2) for the MM1 simulations and methane(1), methane(2) and the 399 intermediate solutes for the MM400 simulations.

The results are given as the RMS deviation of the 18 estimates around three values: (1) around the mean of the 18 estimates; (2) around the best estimate (found by combining the data from all 18 simulations and doing one large calculation); and (3) around the correct answer (in this case, zero). The results are shown in Table VI.

The ME1 and ME400 simulation sets were used to estimate the RMS error in the calculated solvation free energy difference between methane and ethane (see Fig. 3). The ME1 simulations were done analogously to MM1, and the ME400 simulations were done analogously to MM400. The results were calculated in the same manner as the results from the methane-to-methane simulations. The correct answer for this calculation is not known, so we will treat the “best” result from the traditional method as correct; this is the FEPc method with ME400 simulations, giving 0.24 kcal/mol. The results are shown in Table VII.

Relative solvation free energies for nine additional solutes were calculated using only the H7 simulations just described; results are shown in Table VIII. This was done using the CRS method [eqs. (14) and (15)], where the composite reference state consisted of only the coordinates from the “nothing,” methane, ethane, propane, X, X_2 , and X_3 simulations as above; the perturbed states were not included in the composite reference state. The coordinates were modified such that they would correspond to the correct bond-lengths of the perturbed states where necessary. Parameters for methanol were taken from Sun and Kollman.²⁰ Parameters for dimethylether were developed analogously to those for methanol. Parameters for methylamine and dimethylamine are from Morgantini and Kollman.²¹ Meng et al.²² give parameters for methylammonium. The fluorocarbon parameters are from Gough et al.²³ (see Tables I, II, and III for parameters).

Results and Discussion

In this study, the expected errors for different free energy calculation methods are compared from both a theoretical and an empirical perspective. To do the theoretical comparison it was necessary to assume that the estimates were based on a large number of independent samples. Assuming this was valid, it was shown that the IPS method has a

TABLE IV.
Relative Free Energies Calculated Using the FEP, IPS, and CRS Methods (Based on H7 Simulations).^a

Reference state	Nothing	Methane	Perturbed state				
			Ethane	Propane	X	X ₂	X ₃
Nothing	0.00	1.99	4.66	146	2.03	2.16	4.25
		1.91	1.39	DNC ^b	2.20	2.40	2.98
		2.71	2.95	3.11	2.19	2.52	2.84
Methane	-1.91	0.00	0.14	0.50	0.71	0.90	0.91
			0.18	0.05	-0.44	-0.18	0.11
			0.24	0.40	-0.52	-0.19	0.13
Ethane	-1.39	-0.18	0.00	0.27	1.83	0.46	0.54
				0.17	-0.68	-0.48	-0.18
				0.16	-0.76	-0.43	-0.11
Propane	DNC ^b	-0.05	-0.17	0.00	4.46	3.32	2.17
					-1.32	-0.47	-0.24
					-0.92	-0.59	-0.26
X	-2.20	0.55	0.49	6.95	0.00	0.33	0.77
						0.35	0.71
						0.33	0.66
X ₂	-2.52	0.63	0.62	0.80	-0.35	0.00	0.34
							0.34
							0.32
X ₃	-2.84	0.50	0.32	0.32	-0.71	-0.34	0.00

^aThe row of each cell indicates the reference state used for the calculation; the column indicates the perturbed state. The first number in each cell is the result using the FEP method, the second is for the IPS (self consistent) method, and the last is for the CRS method (done by combining simulations of all solutes to form the composite reference state).

^bDid not converge.

smaller expected RMS error than the FEP method. It was also shown that to keep the errors in the estimates small, accessible conformations of the perturbed state must also be accessible during a simulation of the reference state; this was used to

argue that the CRS method should also reduce the magnitude of the expected errors.

Unfortunately, an interpretation of the empirical comparison of the methods is not straightforward, in part, because only a small number of systems

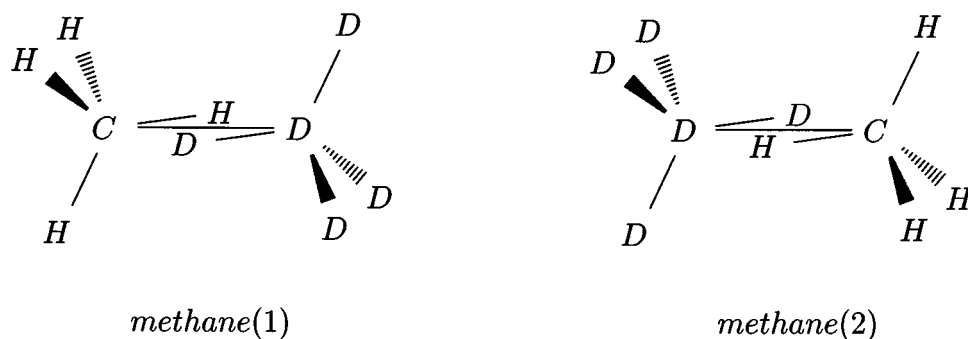


FIGURE 2. Topology for MM solutes. “C” indicates a carbon, “H” indicates a hydrogen, and “D” indicates a dummy atom (no charge or van der Waals interaction). See Table III for van der Waals parameters (atom type CT and HC). Charges are: $q_C = -0.464$, $q_H = 0.116$.

TABLE V.
Previously Calculated Relative Solvation
Free Energy for Selected Hydrocarbons.²⁰

Perturbation	Previous calculated free energy ^a (kcal / mol)
Nothing → Methane	2.67 ± 0.04
Methane → Ethane	0.15 ± 0.07
Ethane → Propane	0.18 ± 0.09

^aSee Sun et al.²⁰

were examined. Additionally, there are many different ways to perform the calculations consistently with each method, and, again, only a small number of them could be examined. For example, to ensure reasonable estimates we performed long (400-ps) simulations for each of the 18 methane-to-methane and the 18 methane-to-ethane simulations. We assumed that conclusions drawn from these simulations would be appropriate for shorter simulations that are more typical. With these caveats in mind, results from the five sets of simulations performed will be described and compared.

Relative solvation free energies for the H7 solutes are shown in Table IV. The correct relative solvation free energy for the parameters used is not known so we will compare the solvation free energies of methane, ethane, and propane with results from a simulation reported by Sun and Kollman²⁰ (see Table V). The RMS errors in the free energy estimates were not determined for these calculations; we only show whether or not the estimates are in reasonable agreement with previous calculations.

Results using the FEP method (first number in each cell) depend on the direction of the perturbation. For example, the relative solvation free energy for methane to ethane (using the methane simulation) is calculated as 0.14 kcal/mol and the relative solvation free energy for ethane to propane (using the ethane simulations) is 0.27 kcal/mol. Both results are in good agreement with the previously calculated values given in Table V. Even the “nothing” to methane perturbation is surprisingly accurate if the calculation is done using the X atom as an intermediate (“nothing” to X gives 2.03 kcal/mol and X to methane gives 0.55 kcal/mol) resulting in a combined value of 2.58 kcal/mol. However, the reverse calculations (where the calculation is based on simulations of the larger of the two solutes) give values in error by over 1 kcal/mol for the propane to ethane and ethane to methane calculations, and the error in the methane to “nothing” calculation is even larger.

These results are consistent with our expectations in that accessible conformations of the perturbed state must be accessible during a simulation of the reference state to give good estimates. For example, adding a methyl group to methane can be permissible because the simulation of methane will occasionally generate water structures appropriate for ethane. Conversely, removing a methyl group during a simulation of ethane does not give good estimates because the simulation of ethane will not generate water structures appropriate for methane. A slightly oversimplified description of this situation is that the water structure around the larger solute is a subset of the

TABLE VI.
Relative Free Energy Estimates for Methane to Methane.^a

Method	Relative free energy (kcal / mol)					
	Mean ± RMS deviation		Best ± RMS deviation		0.000 ± RMS deviation	
	MM1	MM400	MM1	MM400	MM1	MM400
TI	−18.443 ± 27.835	−0.096 ± 0.458	−18.443 ± 27.835	−0.096 ± 0.458	0.000 ± 33.390	0.000 ± 0.468
FEPf	1.620 ± 0.508	0.085 ± 0.208	1.274 ± 0.615	0.029 ± 0.215	0.000 ± 1.698	0.000 ± 0.224
FEP _r	−1.774 ± 0.268	−0.089 ± 0.217	−1.707 ± 0.277	−0.039 ± 0.222	0.000 ± 1.794	0.000 ± 0.234
FEP _c	−0.077 ± 0.323	−0.002 ± 0.194	−0.217 ± 0.352	−0.005 ± 0.194	0.000 ± 0.332	0.000 ± 0.194
IPS ₀	−0.041 ± 0.168	−0.004 ± 0.229	−0.039 ± 0.168	−0.007 ± 0.230	0.000 ± 0.173	0.000 ± 0.230
IPS _{SC}	−0.040 ± 0.163	−0.009 ± 0.225	−0.038 ± 0.163	−0.010 ± 0.225	0.000 ± 0.168	0.000 ± 0.225
CRS	−0.040 ± 0.163	−0.002 ± 0.247	−0.038 ± 0.163	−0.002 ± 0.247	0.000 ± 0.168	0.000 ± 0.247

^aCalculations were done using two sets of 18 methane-to-methane simulations, each 401 ps long (the MM1 and MM400 simulations). Root mean square (RMS) deviations are of the individual simulations around the estimate given. They are not the standard deviation of the estimate.

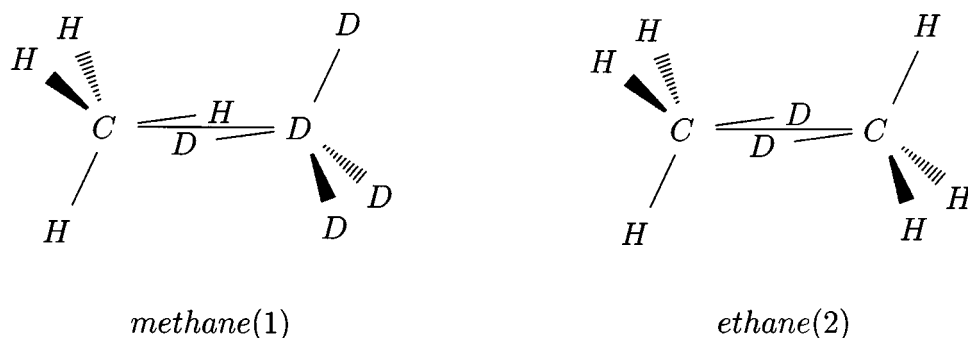


FIGURE 3. Topology for ME solutes. “C” indicates a carbon, “H” indicates a hydrogen, and “D” indicates a dummy atom (no charge or van der Waals interaction). See Table III for van der Waals parameters (atom type CT and HC). Charges for methane are: $q_C = -0.464$, $q_H = 0.116$; charges for ethane are: $q_C = -0.027$, $q_H = 0.009$.

water structure around the smaller solute, allowing simulations of the smaller solute to adequately sample water structure of both solutes.

The IPS method (second number in each cell) gives results in good agreement with those given in Table V for changes in the solute involving the addition of one methyl group (or less). For larger changes in the solute the IPS method does not give good estimates. Note that the IPS calculations require simulations of both solutes, so the two IPS entries for any pair of solutes are identical (unlike the FEP calculation, which is based on simulations of only one of the two solutes). The CRS method (last number in each cell), which uses a single composite reference state (formed by combining simulations of all seven solutes) to calculate the relative free energies, gives results in good agreement with those given in Table V for all pairs of solutes.

A problem with the comparison of results from the H7 simulations is that it shows only whether or not the results are in reasonable agreement with previous calculations. To get reasonable estimates of the errors for each of the different methods, we have examined some simple perturbations in much greater detail.

Results from the 400-window methane-to-methane (see Fig. 2) simulations are given in Table VI (the MM400 simulations). The correct relative solution free energy is obviously zero, so the last two columns are the most useful for comparing RMS deviations of the estimates for the different methods. This shows that the FEP, IPS, and CRS methods all give about the same RMS error (about 0.2 kcal/mol), and the TI method gives an RMS error about twice as large. In general, this is consistent with our expectations because the theoretical advantage of the IPS and CRS methods is that

TABLE VII.
Relative Free Energy Estimates for Methane to Ethane.^a

Method	Relative free energy (kcal / mol)					
	Mean \pm RMS deviation		Best \pm RMS deviation		0.240 \pm RMS deviation	
	ME1	ME400	ME1	ME400	ME1	ME400
TI	217.664 \pm 14.914	0.520 \pm 0.489	217.664 \pm 14.914	0.520 \pm 0.489	0.240 \pm 217.935	0.240 \pm 0.564
FEPf	0.181 \pm 0.080	0.257 \pm 0.210	0.175 \pm 0.080	0.214 \pm 0.214	0.240 \pm 0.100	0.240 \pm 0.211
FEP _r	-1.199 \pm 0.382	0.204 \pm 0.210	-0.964 \pm 0.449	0.265 \pm 0.219	0.240 \pm 1.489	0.240 \pm 0.213
FEP _c	-0.509 \pm 0.189	0.230 \pm 0.187	-0.394 \pm 0.222	0.240 \pm 0.187	0.240 \pm 0.773	0.240 \pm 0.187
IPS ₀	0.174 \pm 0.074	0.252 \pm 0.183	0.171 \pm 0.074	0.230 \pm 0.184	0.240 \pm 0.100	0.240 \pm 0.183
IPS _{SC}	0.172 \pm 0.074	0.249 \pm 0.174	0.169 \pm 0.074	0.235 \pm 0.174	0.240 \pm 0.100	0.240 \pm 0.174
CRS	0.172 \pm 0.074	0.237 \pm 0.184	0.169 \pm 0.074	0.222 \pm 0.185	0.240 \pm 0.100	0.240 \pm 0.185

^aCalculations were done using two sets of 18 methane-to-ethane simulations, each 401 ps long (the ME1 and ME400 simulations). Root mean square (RMS) deviations are of the individual simulations around the estimate given. They are not the standard deviation of the estimate.

TABLE VIII.
Relative Solvation Free Energy Calculated Using the CRS Method [Eqs. (14) and (15)], Where the Composite Reference State Consists of only Nothing, Methane, Ethane, Propane, X, X₂, and X₃ (See Text). These Results Were Calculated Based on the Composite Reference State Formed by Combining the H7 Simulations.

Perturbation	Free energy (kcal / mol)			
	CRS	Rank ^a	Previous calc.	Rank ^a
Ethane → CH ₃ OH-1 ^b	-6.11	—	-6.89 ^d	—
Ethane → CH ₃ OH-2 ^b	-5.55	—	-6.89 ^d	—
Ethane → CH ₃ OH-3 ^b	-5.43	—	-6.89 ^d	—
Ethane → CH ₃ OH ^c	-5.77	2	-6.89 ^d	2
Propane → O(CH ₃) ₂	-3.02	5	-3.5 ^e	5
Nothing → CH ₃ NH ₂ -1 ^b	-3.05	—	-3.57 ^f	—
Nothing → CH ₃ NH ₂ -2 ^b	-2.12	—	-3.57 ^f	—
Nothing → CH ₃ NH ₂ -3 ^b	-1.98	—	-3.57 ^f	—
Nothing → CH ₃ NH ₂ ^c	-2.58	3	-3.57 ^f	3
Nothing → CH ₃ NH ₃ ⁺	-58.9	1	-66.9 ^g	1
Nothing → (CH ₃) ₂ NH-1 ^b	-0.35	—	-1.95 ^f	—
Nothing → (CH ₃) ₂ NH-2 ^b	-0.03	—	-1.95 ^f	—
Nothing → (CH ₃) ₂ NH ^c	-0.21	4	-1.95 ^f	4
Methane → CH ₃ F	-2.44	6	-2.3 ^h	7
Methane → CH ₂ F ₂	-2.21	7	-2.7 ⁱ	6
Methane → CHF ₃	-1.89	8	-1.7 ⁱ	8
Methane → CH ₄	0.42	9	0.44 ^h	9

^aRanked based on solvation free energy, using solvation free energies for methane, ethane, and propane from Table V.
^bOnly one of the possible hydrogen orientations was used.
^cAll the possible hydrogen orientations were used.
^dSee Sun and Kollman.²⁰
^eUsed different method.³²
^fSee Morgantini and Kollman.²¹
^gSee Meng et al.²² (this does not include the Born cutoff correction of -20.5 kcal / mol).
^hSee Gough et al.²³
ⁱAverage of two indirect paths.²³

they required only *some* overlap between states. In this case, each calculation is divided into many small calculations (using 399 intermediate states), so the overlap between adjacent states is apparently adequate to allow each simulation to sample enough of the perturbed state to get good free energy estimates (even for the FEP method). Thus poor overlap is not the limiting factor in these calculations.

We were moderately surprised that the TI errors were not more similar to the FEP errors. This suggests that the rate of change in free energy with respect to the coupling parameter (λ) is not constant over each window [see eq. (2) and associated text] for this calculation. In fairness to the TI method, it should be noted that no effort was made to choose intermediates such that the free energy change would be more nearly linear with λ .

To evaluate how well the methods would perform for calculations involving poor overlap between the two states of interest, the relative free energy was calculated using no intermediate states (the MM1 simulations). The results are consistent with our expectations in that the IPS and CRS methods give the smallest RMS errors (about 0.2 kcal/mol), followed by the FEP methods (from 0.3 to 1.8 kcal/mol). The TI method gives the worst results, as expected, because an assumption used to find eq. (2) is not appropriate here. The most surprising result from this set of calculations is that RMS errors using the combined FEP method (FEPc) are only about one fifth the size of the RMS errors using the forward or the reverse FEP methods (0.3 kcal/mol vs. 1.7 and 1.8 kcal/mol). This seemed a bit fortuitous considering that each FEPc estimate is calculated by averaging an FEPf and

and FEPr estimate. A possible explanation for this is that the symmetry of this calculation (methane to methane) led to cancelation of errors that would not be expected for most useful calculations. If the FEPc results are ignored (we will come back to this point later), these results show that, for calculations involving poor overlap between the states of interest, the IPS and CRS methods are clearly superior to the FEP and TI methods. This is consistent with results from previous comparisons of the FEP and IPS methods done by Saito and Nakamura²⁴ and Ferguson.²⁵

The real question we hoped to answer with these simulations was whether it is better to use many windows (giving better overlap between states) or few windows (allowing more simulation of the states of interest). With this in mind, the MM1 and MM400 simulations were performed such that they used approximately the same amount of CPU time. Our results show that the MM1 simulations give smaller RMS errors (for the IPS and CRS methods), but the difference is small enough that it may not be meaningful. In fact, the MM1 simulations generated twice as much useful data as the MM400 simulations (because half the MM400 simulation time was devoted to equilibrating the system, which was not required for the MM1 simulations), possibly explaining why the ratio of the MM400 and MM1 errors is approximately equal to the square root of 2. It is also possible that a different choice of intermediate states would improve the results from the MM400 simulations. Thus, the results of the MM1 and MM400 simulations do not indicate whether it is better to use 1 or 400 windows. Obviously, an intermediate number of windows may be better for this case, but we chose not to pursue the matter further, in large part because the optimal number of windows is likely to be system dependent.

The major disadvantage of using the methane-to-methane simulations to compare free energy calculation methods is that the symmetry of this change may lead to cancelation of errors that would not take place with most relevant systems. We therefore performed and examined a methane-to-ethane calculation in a manner similar to the methane-to-methane calculation described previously. Unfortunately, the correct relative solvation free energy for these parameters is not known (results shown in Table V are not adequately converged for this comparison). So we will use the estimate from the traditional method (0.24 kcal/mol) using 400 windows and the FEPc method (see Table VII). Because this may not be

the correct result, we are biasing conclusions we draw regarding the errors in the different methods in favor of this method.

Errors for the different methods applied to the 400 window methane-to-ethane simulations are given in Table VII (the ME400 simulations). As with the MM400 simulations, the FEP, IPS, and CRS methods all give about the same RMS error (about 0.2 kcal/mol), so poor overlap is not the limiting factor in these calculations. The TI method gives RMS errors, about three times larger, so the free energy is not nearly linear over the $\Delta\lambda$ used here.

When one window is used (ME1 simulations), the IPS and CRS methods give the smallest RMS errors (0.1 kcal/mol). Two additional points are worth noting here. First, the FEPc method does not give good estimates (its RMS error is about 0.8 kcal/mol). This shows that, in general, the FEPc method should not be used for calculations involving large perturbations (despite the small error obtained using the FEPc method and the MM1 simulations). The second important point is that the FEPf method gives errors as small as the IPS and CRS methods, whereas the FEPr method gives errors more than an order of magnitude larger. This is consistent with the expectation that simulations of the smaller solute can be used to give estimates of the relative solvation free energy of the larger solute, but the converse is not true. This is also qualitatively consistent with the FEP results from the H7 simulations discussed previously.

A comparison of RMS errors from the ME1 and ME400 simulations shows that the ME1 simulations give smaller errors, but we have the same concerns discussed above in the comparison of the MM1 and MM400 simulations. It should be noted, however, that if a relative free energy of 0.15 kcal/mol (as reported by Sun and Kollman²⁰; see Table V) is assumed to be correct, then free energy estimates based on the one-window simulation are in excellent agreement and, more importantly, the RMS deviation from 0.15 kcal/mol is smaller than the RMS deviation from 0.24 kcal/mol. The converse is true for the 400-window simulations. In summary, our results suggest that a single window gives better results for this case, but the difference does not seem to be significant.

The observation that simulations of methane can be used to estimate relative solvation free energies of ethane suggests the use of simulations of a reference solute to estimate the relative free energy of modified solutes that have not been simulated. This has been done previously. For ex-

ample, free energy changes have been estimated using the rate of change in free energy as a function of atomic parameters, calculated using a single simulation.^{26,27} Also, van Gunsteren and coworkers used "soft" atoms, which allow sampling of conformations corresponding to the presence of a real atom or no atoms at the soft-atom position.²⁸

Instead of using a simulation of one solute, we propose combining coordinates from simulations of a set of solutes using the CRS method. The advantage of doing this is it gives more flexibility for selecting the regions of conformation space that will be sampled. If this is done carefully, a relatively small set of appropriately chosen solutes could be used to sample the important conformations of a larger number of solutes that are not simulated. This can then be used to find the relative free energies of solutes that have not been simulated.

To evaluate this method, we found the composite reference state for the saved coordinates from the H7 simulations, as described previously. Equations (14) and (15) were used to calculate the relative solvation free energy of nine solutes that were not simulated. They are methanol, dimethylether, methylamine, dimethylamine, methylammonium, and four fluoromethanes. These solutes were chosen because they all have structural similarities to at least one of the seven simulated solutes and, additionally, they have had their solvation free energies calculated previously so that comparisons can be made. The results are given in Table VIII.

The estimated solvation free energies for the fluorocarbons are in good agreement with previously calculated results. This shows that heavy atoms can be added to the solute in a single step. The free energies for the hydrogen bonding solutes are all less negative than the previously calculated values, indicating that the water structure for the seven hydrophobic simulated solutes did not sample all of the solvent structure for solutes involved in electrostatic interactions with the solvent. For example, the solvation free energies for methanol and methylamine are both underestimated by about 1 kcal/mol. Although quantitatively disappointing, this is impressive qualitatively, because the calculation shows much of the expected hydrogen bonding between the perturbed solute and water (meaning that the simulations did sample perturbed state water structures), even though it is based on simulations of only hydrophobic solutes.

Even the calculation of methylammonium, which involves adding a net charge, shows much of the expected electrostatic interaction with water.

The least impressive calculation for neutral solutes underestimates the solvation free energy for dimethylamine by almost 2 kcal/mol. We found this surprising because the solvation free energy for dimethylether is underestimated by only 0.5 kcal/mol, even though the two solutes are structurally similar. To understand why this is the case, interactions between dimethylamine and a single water molecule as well as interactions between dimethylether and a single water molecule were examined. This was done by moving a water molecule around each of the two solutes, and identifying the low-energy conformations. For each solute, a total of 500 million conformations were generated. Almost 20,000 conformations within 0.6 kcal/mol of the minimum energy conformation for dimethylether were found, whereas only about 2300 conformations within 0.6 kcal/mol of the minimum energy conformation for dimethylamine were found (this order of magnitude difference was constant for any energy difference up to at least 3 kcal/mol).

Examination of these conformations shows that, for near-optimal water–dimethylether interactions, the water can occupy any position near one, or both, of the lone pairs on dimethylether's oxygen. However, near-optimal interactions with dimethylamine require that the water approach the lone pair of dimethylamine's nitrogen from above the plane of the three heavy atoms, allowing the water to occupy only one end of the region occupied by waters interacting with dimethylether. This shows that it is easier to find good interactions between dimethylether and a single water than between dimethylamine and a single water. We assume this is also true for bulk water, accounting for the difference in errors. We also assume that the error would be reduced if a better choice of reference solutes is used to define the composite reference state, or longer simulations are performed.

The solvation free energies of solutes found using this method clearly are not as accurate as calculations where all solutes are simulated; however, the ranking of absolute solvation free energy, based on previously calculated values, is almost identical to the ranking based on the CRS solvation free energies (see Table VIII). This is of interest because it suggests that this method could be used to estimate and rank interaction free energies between a receptor and a large number of proposed ligands, without simulating each of them. A sce-

nario that might be useful for doing structure-based ligand design with biological macromolecules starts by simulating the macromolecule bound to a known ligand and calculating the rate of change in the free energy with respect to parameters (see Cieplak et al.^{27,29} for an example of the use of free energy derivatives). This, in conjunction with examination of the structures using molecular graphics, could be used to suggest ligands that might bind the macromolecules more tightly. These proposed ligands would then be quickly ranked, based on their estimated binding free energies, using the method described here. The most promising ligands could then be examined carefully using more accurate free energy calculation methods, or their binding could be determined experimentally.

Conclusions

In the "Theoretical Background" section, we argue that the IPS and CRS methods should give smaller expected errors than the FEP and TI methods. An empirical comparison of the methods shows that, in general, the IPS and CRS methods give better results if the overlap between the states is small (as for the MM1 and ME1 simulations). However, if the overlap is very good (as for the MM400 and ME400 simulations) then the FEP, IPS, and CRS methods give about the same RMS error. Our results also suggest that using few intermediate states (and the IPS or CRS method) may give better results than using many states, but the difference does not appear to be significant. Considering that the IPS and CRS methods never give significantly larger errors, our results suggest that the IPS or CRS methods should be used in place of the FEP or TI methods, particularly if the overlap between states is not known to be good (it should be kept in mind that the overlap may be inadequate for even the IPS or CRS methods, in which case additional intermediate states would be needed, regardless of the method used).

We also show that estimating the solvation free energies of solutes that are not simulated is possible, although the results are less accurate. For the solutes examined, the ranking of solvation free energy is almost identical to the previously calculated ranking. This may be useful for estimating the relative free energies of a large number of compounds to identify which would be worth examining further. We are currently exploring the possibility of using this procedure to estimate the

relative binding free energy for ligand/macro-molecule interactions so that we can identify ligand modifications worthy of more accurate free energy calculations. We anticipate that this will be useful for improving the binding of drugs or lead compounds to biomolecules.

Acknowledgments

R. J. R. gratefully acknowledges support from NIH Grant GM-39552 (G. L. Kenyon); P. A. K. thanks the NSF (Grant CHE-94-13472).

Appendix

RMS ERRORS IN FREE ENERGY ESTIMATES AND OPTIMAL INTERMEDIATE STATES

The RMS Error in Free Energy Estimates for Simulations Performed on State 0 and Perturbed to State 1

The RMS error of a free energy estimate is:

$$RMS(\Delta g_{10} - \Delta G_{10}) = \left\{ E \left[(\Delta g_{10} - \Delta G_{10})^2 \right] \right\}^{\frac{1}{2}} \quad (17)$$

where ΔG_{10} is the free energy, Δg_{10} is an estimate of the free energy, and $E(\dots)$ indicates the expectation value calculated over all possible sets of samples. The error can be written as:

$$\Delta g_{10} - \Delta G_{10} = -RT \ln \frac{1}{n_0} \sum_{i=1}^{n_0} \frac{\rho_1(\mathbf{r}_i)}{\rho_0(\mathbf{r}_i)} \quad (18)$$

where coordinates, \mathbf{r}_i , are chosen with a Boltzmann probability, $\rho_0(\mathbf{r}_i)$. Substituting eq. (18) into eq. (17), and finding the expectation value, gives:

$$\begin{aligned} RMS(\Delta g_{10} - \Delta G_{10}) &= RT \left[\int_{\mathbf{r}_{n_1}} \cdots \int_{\mathbf{r}_1} \left(\ln \frac{1}{n_0} \sum_{i=1}^{n_0} \frac{\rho_1(\mathbf{r}_i)}{\rho_0(\mathbf{r}_i)} \right)^2 \right. \\ &\quad \left. \times \prod_{k=1}^{n_1} \rho_1(\mathbf{r}_k) d\mathbf{r}_k \right]^{\frac{1}{2}} \quad (19) \end{aligned}$$

We cannot solve this exactly, so the logarithm is approximated with the first term of its Taylor series expansion:

$$\ln \frac{1}{n_0} \sum_{i=1}^{n_0} \frac{\rho_1(\mathbf{r}_i)}{\rho_0(\mathbf{r}_i)} \approx \frac{1}{n_0} \sum_{i=1}^{n_0} \frac{\rho_1(\mathbf{r}_i)}{\rho_0(\mathbf{r}_i)} - 1 \quad (20)$$

For sets of samples that give good estimates of the free energy ($\Delta g_{10} - \Delta G_{10} \approx 0$), this approximation is obviously reasonable [see eq. (18)]. The approximation is not good for sets of samples that overestimate or underestimate the free energy. Problems resulting from this approximation may be most serious as $\frac{1}{n_0} \sum_{i=1}^{n_0} \frac{\rho_1(\mathbf{r}_i)}{\rho_0(\mathbf{r}_i)}$ approaches zero because the right-hand side of eq. (20) approaches negative unity instead of negative infinity. This situation is most likely to occur if there are a significant number of possible set of samples that sample the perturbed state only in inaccessible regions. For example, if most of the conformations sampled during simulations of the reference state are not accessible to the perturbed state, the estimate of the RMS error given in this section is likely to underestimate the correct error.

Accepting these problems, the estimate of the RMS error can be simplified by substituting eq. (20) into eq. (19). Expanding the resulting integrand, integrating the cross terms, and moving the summation out from under the integral gives:

$$RMS(\Delta g_{10} - \Delta G_{10}) \approx RT \left(\frac{1}{n_0} \int \frac{\rho_1^2(\mathbf{r})}{\rho_0(\mathbf{r})} d\mathbf{r} - \frac{1}{n_0} \right)^{\frac{1}{2}} \quad (21)$$

RMS Error in Free Energy Estimates and the Optimal Perturbed State for Simulations Performed on States 0 and 1, Perturbed to State b

If sampling on states 0 and 1 is independent, then:

$$RMS^2(\Delta g_{10} - \Delta G_{10}) = RMS^2(\Delta g_{b0} - \Delta G_{b0}) + RMS^2(\Delta g_{b1} - \Delta G_{b1})$$

Sampling on states 0 and 1 is independent, so [see eq. (21)]:

$$RMS^2(\Delta g_{10} - \Delta G_{10}) \approx (RT)^2 \left(\int \frac{\rho_b^2(\mathbf{r})}{n_0 \rho_0(\mathbf{r})} d\mathbf{r} - \frac{1}{n_0} + \int \frac{\rho_b^2(\mathbf{r})}{n_1 \rho_1(\mathbf{r})} d\mathbf{r} - \frac{1}{n_1} \right)$$

Using Euler's equation (see Boas,³⁰ Pierre,³¹ or any other description of variational calculus), the

$\rho_{b,opt}$ that minimizes $RMS^2(\Delta g_{10} - \Delta G_{10})$ can be found:

$$\frac{\partial \left[\frac{\rho_{b,opt}^2(\mathbf{r})}{n_0 \rho_0(\mathbf{r})} + \frac{\rho_{b,opt}^2(\mathbf{r})}{n_1 \rho_1(\mathbf{r})} - 2\lambda_b \rho_{b,opt}(\mathbf{r}) \right]}{\partial \rho_{b,opt}(\mathbf{r})} = 0$$

Differentiating and solving for $\rho_{b,opt}(\mathbf{r})$ gives:

$$\rho_{b,opt}(\mathbf{r}) = \lambda_b \frac{n_0 \rho_0(\mathbf{r}) n_1 \rho_1(\mathbf{r})}{n_0 \rho_0(\mathbf{r}) + n_1 \rho_1(\mathbf{r})}$$

where λ_b is found to normalize $\rho_{b,opt}$.

RMS Error in Free Energy Estimates and the Optimal Reference State for Simulations Performed on State s, Perturbed to States 0 and 1

The error in the free energy estimate is:

$$\Delta g_{10} - \Delta G_{10} = -RT \ln \frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\rho_0(\mathbf{r}_i)}{\rho_s(\mathbf{r}_i)} + RT \ln \frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\rho_1(\mathbf{r}_i)}{\rho_s(\mathbf{r}_i)}$$

Using the first term of the logarithm's Taylor series expansion to approximate the logarithm [eq. (20)] gives:

$$\begin{aligned} \Delta g_{10} - \Delta G_{10} &\approx -RT \left(\frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\rho_0(\mathbf{r}_i)}{\rho_s(\mathbf{r}_i)} - 1 \right) \\ &\quad + RT \left(\frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\rho_1(\mathbf{r}_i)}{\rho_s(\mathbf{r}_i)} - 1 \right) \\ \Delta g_{10} - \Delta G_{10} &\approx -RT \left(\frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\rho_0(\mathbf{r}_i) - \rho_1(\mathbf{r}_i)}{\rho_s(\mathbf{r}_i)} \right) \end{aligned}$$

This gives an RMS error of:

$$\begin{aligned} RMS(\Delta g_{10} - \Delta G_{10}) &\approx RT \left[\int_{\mathbf{r}_{n_s}} \dots \int_{\mathbf{r}_1} \left(\frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\rho_0(\mathbf{r}_i) - \rho_1(\mathbf{r}_i)}{\rho_s(\mathbf{r}_i)} \right)^2 \right. \\ &\quad \left. \times \prod_{k=1}^{n_s} \rho_s(\mathbf{r}_k) d\mathbf{r}_k \right]^{\frac{1}{2}} \end{aligned}$$

Expanding the integrand and integrating where possible gives:

$$\text{RMS}(\Delta g_{10} - \Delta G_{10})$$

$$\approx RT \left(\frac{1}{n_s} \int \frac{[\rho_0(\mathbf{r}) - \rho_1(\mathbf{r})]^2}{\rho_s(\mathbf{r})} d\mathbf{r} \right)^{\frac{1}{2}} \quad (22)$$

Using Euler's equation as above, the $\rho_{s,opt}$ that minimizes this equation can be found:

$$\frac{\partial \left[\frac{[\rho_0(\mathbf{r}) - \rho_1(\mathbf{r})]^2}{\rho_{s,opt}(\mathbf{r})} + \lambda_s^{-2} \rho_{s,opt}(\mathbf{r}) \right]}{\partial \rho_{s,opt}(\mathbf{r})} = 0$$

Differentiating and solving for $\rho_{s,opt}(\mathbf{r})$ gives:

$$\rho_{s,opt}(\mathbf{r}) = \lambda_s |\rho_0(\mathbf{r}) - \rho_1(\mathbf{r})|$$

where λ_s is found to normalize ρ_s .

Unfortunately, there is a significant difficulty with this optimal state; it allows the RMS error in the two intermediate estimates, Δg_{0s} and Δg_{1s} , to approach infinity, making them impossible to calculate in practice. We are not able to determine if this is a result of approximating the logarithm with part of its Taylor series expansion, or if it would be the case even if the RMS error could have been calculated without making any approximations.

References

1. C. H. Bennett, *J. Comput. Phys.*, **22**, 245–268 (1976).
2. S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, *J. Comput. Chem.*, **13**, 1011–1021 (1992).
3. J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, New York, 1987.
4. D. L. Beveridge and F. M. DiCapua, *Annu. Rev. Biophys. Biophys. Chem.*, **18**, 431–492 (1989).
5. T. P. Straatsma and J. A. McCammon, *Annu. Rev. Phys. Chem.*, **43**, 407–435 (1992).
6. P. A. Kollman, *Chem. Rev.*, **93**, 2395–2417 (1993).
7. D. A. Pearlman, *J. Phys. Chem.*, **98**, 1487–1493 (1994).
8. R. W. Zwanzig, *J. Chem. Phys.*, **22**, 1420–1426 (1954).
9. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, New York, 1987.
10. W. F. van Gunsteren, *Prot. Eng.*, **2**, 5–13 (1988).
11. B. L. Tembe and J. A. McCammon, *Comput. Chem.*, **8**, 281–283 (1984).
12. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, Cambridge University Press, New York, 1989.
13. G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, **23**, 187–199 (1977).
14. D. A. Pearlman, D. A. Case, J. W. Caldwell, G. L. Seibel, U. C. Singh, P. Weiner, and P. A. Kollman, University of California at San Francisco, 1991.
15. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.*, **79**, 926–935 (1983).
16. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.*, **81**, 3684–3690 (1984).
17. W. F. van Gunsteren and H. J. C. Berendsen, *Mol. Phys.*, **34**, 1311–1327 (1977).
18. S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case, *J. Comput. Chem.*, **7**, 230–252 (1986).
19. W. L. Jorgensen and C. Ravimohan, *J. Chem. Phys.*, **83**, 3050–3054 (1985).
20. Y. Sun and P. A. Kollman, *J. Comput. Chem.*, **16**, 1164–1169 (1995).
21. P.-Y. Morgantini and P. A. Kollman, *J. Am. Chem. Soc.*, **117**, 6057–6063, 1995.
22. E. C. Meng, P. Cieplak, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.*, **116**, 12061–12062 (1994).
23. C. A. Gough, D. A. Pearlman, and P. A. Kollman, *J. Chem. Phys.*, **99**, 9103–9110, 1993.
24. M. Saito and H. Nakamura, *J. Comput. Chem.*, **11**, 76–81 (1990).
25. D. M. Ferguson, *J. Chem. Phys.*, **99**, 10086–10087 (1993).
26. P. E. Smith and W. F. van Gunsteren, *J. Chem. Phys.*, **100**, 577–585 (1994).
27. P. Cieplak, D. A. Pearlman, and P. A. Kollman, *J. Chem. Phys.*, **101**, 627–633 (1994).
28. T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, and W. F. van Gunsteren, *Chem. Phys. Lett.*, **222**, 529–539 (1994).
29. P. Cieplak and P. A. Kollman, *J. Comput.-Aided Mol. Des.*, **7**, 291–304 (1993).
30. M. L. Boas, *Mathematical Methods in the Physical Sciences*, 2nd Ed., Wiley, New York, 1983.
31. D. A. Pierre, *Optimization Theory with Applications*, Dover, New York, 1986.
32. D. L. Veenstra, personal communication.